LOCAL-VS-GLOBAL COMBINATORICS

ASAF SHAPIRA

ABSTRACT

Many of the most outstanding open problems in combinatorics relate the local and global properties of large discrete structures. The research aimed at solving these questions led to some of the most important developments in this area, as well as in related areas such as theoretical computer science, additive number theory, and harmonic analysis. In this paper we discuss some of these advances and mention several open problems.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 05C35; Secondary 05D99, 05C65, 05C85, 68R10, 68W20

KEYWORDS

Regularity lemma, quasi randomness, removal lemma, property testing, graph theory, hypergraphs

INTERNATIONAL CONGRESS
OF MATHEMATICIANS
2022 JULY 6–14

Proc. Int. Cong. Math. 2022, Vol. 6, pp. 4682–4708
DOI 10.4171/ICM2022/36

1. INTRODUCTION

Extremal combinatorics is one of the fastest growing areas of research within discrete mathematics. Questions in this area deal with the asymptotic relations between various parameters of large discrete structures such as graphs, hypegraphs, permutations, sets of integers, etc. This area has grown tremendously in the past few decades, both in depth and in breadth, and supplied many spectacular results that affected various other areas of mathematics, such as number theory, group theory, probability theory, information theory, harmonic analysis, and theoretical computer science. Many key insights that were developed in order to solve some of the core problems in extremal combinatorics were later exported to other areas. Perhaps the prime example is Szemerédi's theorem [94], stating that dense sets of integers contain arbitrarily long arithmetic progressions. This theorem motivated some of the most important investigations in extremal combinatorics such as the regularity method in graphs [95] and hypergraphs [54, 73, 81], the theory of quasirandom graphs [23, 98], and the theory of graph limits [68]. Szemerédi's theorem also motivated the development of tools in other areas such as ergodic theory (the multiple recurrence theorem [39, 40]), harmonic analysis (the Gowers norms [53]), number theory (the Green–Tao theorem [57]), and theoretical computer science (the PCP theorem [13, 14] and property testing [48]). See [96] for a more detailed discussion.

In this paper we describe a variety of results and open problems in extremal combinatorics relating local and global properties of graphs and hypergraphs. The first set of problems is related to one of the most influential open problems in extremal combinatorics. To state it, we need the following definitions. An r-graph $\mathcal{H} = (V, E)$ consists of a ground set V (the vertices) and a collection of subsets E (the edges) where each edge in E contains r distinct vertices of V. When $r = 2$, we will use the term *graph* and denote graphs by G. An r-graph $\mathcal{H}$ is *linear* if every pair of vertices $u, v \in V$ belongs to at most one edge of E. A (v, e)-*configuration* in $\mathcal{H}$ is a set of e edges whose union contains at most v vertices. The following conjecture was raised 50 years ago by Brown, Erdős, and Sós [21, 22]. In the next statement, and in the rest of the paper, we use standard $O/\Omega/\Theta/o$ notation.

Conjecture 1.1 (Brown–Erdős–Sós conjecture). *Fix $e \geq 3$ and suppose $\mathcal{H}$ is an n-vertex linear 3-graph without $(e + 3, e)$-configurations. Then $\mathcal{H}$ has $o(n^2)$ edges.*

Note that as in a (v, e)-configuration we fix the number of edges and only bound the number of vertices, such a configuration is a locally dense subset of $\mathcal{H}$. Since a linear $\mathcal{H}$ clearly has at most n^2 edges, what the above conjecture states is that if $\mathcal{H}$ is locally sparse then it is also globally sparse. It is easy to see that if Conjecture 1.1 holds for linear 3-graphs then it holds for arbitrary 3-graphs.

The second set of problems we cover revolves around the *triangle removal lemma* of Ruzsa and Szemerédi [85], devised for the purpose of proving Conjecture 1.1 for the special case $e = 3$, which is widely considered to be one of the cornerstone results of extremal combinatorics. In what follows, a triangle in a graph $G = (V, E)$ is a triple of vertices u, v, w so that $(u, v), (u, w)(v, w) \in E$. A graph is *triangle-free* if it contains no triangle.

Theorem 1.2 (Triangle removal lemma). *For every $\varepsilon > 0$, there is* $\text{Rem}(\varepsilon)$ *so that if G is an n-vertex graph with the property that one should remove at least εn^2 of its edges in order to make it triangle-free, then G contains at least $n^3/\text{Rem}(\varepsilon)$ triangles.*

Note that if G has $n^3/\text{Rem}(\varepsilon)$ triangles, then a random subset of (about) $\text{Rem}(\varepsilon)$ vertices contains a triangle with probability at least $2/3$. In particular, this means that no matter how large n is, most subsets of vertices of size $\text{Rem}(\varepsilon)$ are not triangle-free. We can thus interpret Theorem 1.2 as stating that if G is globally far from being triangle-free, then it is also locally far from being triangle-free.

The third set of problems we discuss is related to the celebrated *regularity lemma* of Szemerédi [95]. One of the first applications of this lemma was the proof Theorem 1.2 in [85]. Since then it has become one of the most important tools for solving extremal problems in graph theory (see [80]). To state it we need a few definitions. Suppose G is a graph and A, B are two disjoint subsets of V. We use $e(A, B)$ to denote the number of edges of E that connect a vertex in A with a vertex in B. We also let $d(A, B) = e(A, B)/|A||B|$ denote the *edge density* between A, B. Finally, we say that the pair (A, B) is ε-regular if $|d(A, B) - d(A', B')| \leq \varepsilon$ for every pair $A' \subseteq A$, $B' \subseteq B$ satisfying $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$. A partition $V_1, \ldots, V_k$ of the vertices of G into k sets is called ε-regular if all but εk^2 of the pairs (V_i, V_j) are ε-regular and all the sets are of equal size n/k (or of sizes $\lfloor n/k \rfloor$ and $\lceil n/k \rceil$). The *order* of such a partition is the number of sets V_i in it (i.e., k above).

Theorem 1.3 (Szemerédi's regularity lemma). *For every $\varepsilon > 0$, there is an $M = M(\varepsilon)$ so that every graph has an ε-regular partition of order k with $1/\varepsilon \leq k \leq M$.*

The rest of the paper is organized as follows. In Section 2 we discuss Conjecture 1.1 and many related questions. In Section 3 we discuss Theorem 1.3 and many of its variants along with their applications. Finally, variants of Theorem 1.2, and their relations to problems in theoretical computer science, are described in Sections 3 and 4. Since it is clearly impossible to cover all themes related to the subject of this paper, or even those related to the above three topics, many important results will be left out.

2. THE BROWN–ERDŐS–SÓS CONJECTURE

In this section we describe several results and open problems related to Conjecture 1.1. We will henceforth use the acronym BESC. Let $f_r(n, v, e)$ denote the largest number of edges in a linear r-graph on n vertices that contains no (v, e)-configuration. Note that Conjecture 1.1 is the statement that $f_3(n, e+3, e) = o(n^2)$. Despite much effort by many researchers, Conjecture 1.1 is wide open, having only been settled for $e = 3$ by Ruzsa and Szemerédi [85] in what has become known as the $(6, 3)$-theorem. To get some perspective on the significance of this special case of Conjecture 1.1, let us just mention that besides its relation to Theorems 1.2 and 1.3 mentioned above, the $(6, 3)$-theorem implies Roth's theorem [82] on 3-term arithmetic progressions in dense sets of integers (see the next subsection). As another indication of the importance of this problem, we note that one of the main driving forces for proving the celebrated hypergraph removal lemma (see Section 3.3) was the hope that it would lead to a proof of Conjecture 1.1.

2.1. Approximate versions of BESC

At present we seem to be quite far from proving Conjecture 1.1. As an indication of the difficulty of Conjecture 1.1 for $e > 3$, let us mention that already the case $e = 4$ (i.e., the statement $f_3(n, 7, 4) = o(n^2)$) implies the notoriously difficult Szemerédi's theorem [93] for 4-term arithmetic progressions, see [28]. It is thus natural to look for approximate versions of this conjecture. Namely, given $e \geq 3$, find the smallest $d = d(e)$ such that $f_3(n, e + d, e) = o(n^2)$. Until very recently, the best result of this type was obtained about 15 years ago by Sárközy and Selkow [87], who proved that

$$f_3\big(n, e + 2 + \lfloor \log_2 e \rfloor, e\big) = o(n^2). \tag{2.1}$$

Since the result of [87], the only advance was obtained by Solymosi and Solymosi [91], who improved the bound for $e = 10$ from $f_3(n, 15, 10) = o(n^2)$ (which follows from (2.1)), to $f_3(n, 14, 10) = o(n^2)$. The first improvement over (2.1) for all large enough e was obtained recently in [26]. Moreover, it shows that one can replace the $\lfloor \log_2 e \rfloor$ "error term" in (2.1) by a much smaller, sublogarithmic, term.

Theorem 2.1. *For every $e \geq 3$,*

$$f_3(n, e + 18 \log e / \log \log e, e) = o(n^2).$$

The main idea of [87] in their proof of (2.1) was the following: The triangle removal lemma actually shows that a linear 3-graph with $\Omega(n^2)$ edges has many $(6, 3)$-configurations. One then defines an auxiliary graph based on these $(6, 3)$-configurations, and uses the triangle removal lemma again in order to double a $(6, 3)$-configuration into a configuration with 7 edges, and so on. The caveat is that each time the number of edges is doubled, the difference between v and e increases by 1, resulting in the $\log_2 e$ error term. The main idea of [91] was to use the 3-*graph removal lemma* (see Theorem 3.8), which is the extension of Theorem 1.2 to 3-graphs, in order to perform a *single* iteration in the style of [87], which instead of doubling the number of edges, (roughly) triples it. The main novelty in [26] is in managing to perform these multiplications an unbounded number of times. To do so, the r-*graph removal lemma* is used (for all uniformities r) in order to sequentially multiply the number of edges by $3, 4, 5, \ldots$. The main difficulty is in ensuring that each time one multiplies a configuration, the difference between v and e increases only by 1. Another challenge is in making sure the configuration has exactly e edges.

Conjecture 1.1 has a more general form (see [87]), stating that for every $2 \leq k < r$ and $e \geq 3$ we have $f_r(n, (r - k)e + k + 1, e) = o(n^k)$. However, as noted in [26], this more general version is, in fact, equivalent to the special case stated as Conjecture 1.1 (corresponding to $k = 2$ and $r = 3$). Moreover, any approximate version of Conjecture 1.1, like that stated in Theorem 2.1, gives analogous approximate versions of the general conjecture.

2.2. Lower bounds for BESC

As we mentioned at the end of the previous subsection, the general form of the BESC states that for every $2 \leq k < r$ and $e \geq 3$ we have $f_r(n, (r-k)e + k + 1, e) = o(n^k)$. It is also widely believed (see [28, 29]) that the following lower bound holds.

Conjecture 2.2. *For every $2 \leq k < r$ and $e \geq 3$, we have $f_r(n, (r-k)e + k + 1, e) > n^{k-o(1)}$.*

Given the difficulty of proving an upper bound for the BESC, one might expect that Conjecture 2.2 would be relatively easy to resolve. As it turns out, this is not the case. Ruzsa and Szemerédi [85] gave an ingenious construction showing that

$$f_3(n, 6, 3) \geq \Omega\big(n \cdot r_3(n)\big) \geq n^{2-o(1)}, \tag{2.2}$$

where $r_3(n)$ is the size of the largest subset of the first n integers without a 3-term arithmetic progression, and the second inequality follows from the well-known construction of Behrend [16] showing that $r_3(n) > n^{1-o(1)}$. Observe that combined with the fact that $f_3(n, 6, 3) = o(n^2)$ mentioned above, this implies Roth's theorem [82] stating that $r_3(n) = o(n)$. This establishes Conjecture 2.2 for $e = 3$, $k = 2$, $r = 3$. Erdős, Frankl, and Rödl [29] later extended this to arbitrary $r \geq 3$ (and $e = 3$, $k = 2$ as in [85]). A result of [8] then verified Conjecture 2.2 for $e = 3$ and arbitrary $2 \leq k < r$.

The key idea in the above results, which handle the case $e = 3$, is to start with a set of integers $X \subseteq [n]$, and construct a Cayley-type r-graph $\mathcal{H}$ in such a way that one can "extract" from any $((r-k)3 + k + 1, 3)$-configuration in $\mathcal{H}$ a nontrivial solution to an equation of the form $ax + by = (a + b)z$ with bounded a, b and $x, y, z \in X$. A simple generalization of [16] shows that there are sets $X \subseteq [n]$ of size $n^{1-o(1)}$ without nontrivial solutions to equations of this type, which can be used to give a bound as in (2.2). The reason why Conjecture 2.2 becomes much harder when $e > 3$ is that when handling more than 3 edges, the linear equation E we can extract from a $((r-k)e + k + 1, e)$-configuration might be one for which there is no $X \subseteq [n]$ of size $n^{1-o(1)}$ without a solution of E. For example, if the equation is $x + y = z + w$ then a set without nontrivial solutions has size $O(\sqrt{n})$.

Let us focus then on the case $e > 3$ and $k = 2$ and $r = 3$. It is easy to check that every $(7, 4)$ or $(8, 5)$-configuration contains a $(6, 3)$-configuration, hence the bounds $f_3(n, 7, 4), f_3(n, 8, 5) \geq n^{2-o(1)}$ follow from (2.2). The situation becomes much harder when $e = 6$, since there is a $(9, 6)$-configuration which does not contain a $(6, 3)$-configuration. Indeed, this is the 3×3 *grid*, denoted $\mathcal{G}_{3\times3}$, namely the 3-graph whose vertices are the nine points in a 3×3 point array, and whose edges correspond to the 6 horizontal and vertical lines of this array. Let $\mathcal{T}$ denote the 3-graph with vertices $1, 2, 3, 4, 5, 6$ and edges $\{1, 2, 3\}, \{3, 4, 5\}, \{5, 6, 1\}$ (this is the unique linear $(6, 3)$-configuration). It is not hard to verify that every linear $(9, 6)$-configuration (in a 3-graph) either contains a copy of $\mathcal{T}$ or is isomorphic to $\mathcal{G}_{3\times3}$. Hence, to prove that $f(n, 9, 6) \geq n^{2-o(1)}$, it would suffice to construct a linear 3-graph with $n^{2-o(1)}$ edges and no copy of either $\mathcal{G}_{3\times3}$ or $\mathcal{T}$.

The above facts led Füredi and Ruszinkó [38] to study various extremal problems related to $\mathcal{G}_{3\times3}$. In particular, they conjectured that there is a $\mathcal{G}_{3\times3}$-free linear 3-graph with $(1/6 - o(1))n^2$ edges. Using a standard probabilistic alterations argument, Füredi and Ruszinkó [38] constructed such a 3-graph with $\Omega(n^{1.8})$ edges. This was slightly improved (as a special case of a more general result) to $\Omega(n^{1.8}\log^{1/5} n)$ in [88]. The following result of [44] makes a significant progress on the conjecture of Füredi and Ruszinkó [38], by improving these results to $\Omega(n^2)$. We, in fact, have the following more general statement.

Theorem 2.3. *For a prime p, and two sets $X, A \subseteq \mathbb{F}_p$, define the following 3-partite 3-graph $\mathcal{H} = \mathcal{H}(X, A)$ on vertex sets V_1, V_2, V_3 where we think of each V_i as a copy of $\mathbb{F}_p$: for every $x \in X$ and $a \in A$, place a 3-edge containing the vertices $x \in V_1$, $x + a \in V_2$ and $x + a^2 \in V_3$ (all operations over $\mathbb{F}_p$). Then, every pair of vertices of $\mathcal{H}$ belongs to at most 2 edges. Also, if there are no $x_1, x_2 \in X$ and $a \in A$ satisfying*

$$4x_1 + 4a \equiv 4x_2 + 1 \pmod{p} \tag{2.3}$$

then $\mathcal{H}$ is $\mathcal{G}_{3\times 3}$-free, and if A has no solution to the equation

$$a + b^2 - b \equiv c^2 \pmod{p} \tag{2.4}$$

in distinct $a, b, c \in A$ then $\mathcal{H}$ is $\mathcal{T}$-free.

It is easy to see that the sets $X = A = \{1, \dots, \lfloor p/8 \rfloor\}$ do not contain a solution to (2.3), hence $\mathcal{H} = \mathcal{H}(X, A)$ has $\Omega(n^2)$ edges and no copy of $\mathcal{G}_{3\times 3}$. Also, since each pair of vertices belongs to at most 2 edges we get that there is also a *linear* $\mathcal{H}$ with the same properties, thus establishing the improved bound for the Füredi–Ruszinkó conjecture stated before Theorem 2.3. The second assertion Theorem 2.3 thus leads to the following problem:

Problem 2.4. Is there $A \subseteq \mathbb{F}_p$ of size $p^{1-o(1)}$ without a solution of (2.4) in distinct a, b, c?

If a set A as above exists, then, to prove that $f(n, 9, 6) \geq n^{2-o(1)}$, one would just need to find $X \subseteq \mathbb{F}_p$ of size $p^{1-o(1)}$ so that A and X have no solution to (2.3). Also, in the spirit of [84], it seems interesting to further study the size of the largest subsets of $\mathbb{F}_p$ without nontrivial solutions to other polynomial equations.

Conjectures 1.1 and 2.2 state nearly matching lower and upper bounds. We conclude this subsection by recalling a problem of Erdős [28], who asked if the exact asymptotic formula $f(n, e + 3, e) = \Theta(n \cdot r_e(n))$ holds, where $r_e(n)$ denotes the size of the largest subset of the first n integers without an e-term arithmetic progression. As of now, the upper bound is not known for any $e \geq 3$, while the lower bound is known only for $e = 3, 4, 5$.

2.3. The Gowers–Long conjecture

The BESC states that a 3-graph without $(e + 3, e)$-configurations has $o(n^2)$ edges. By (2.2), already when $e = 3$, one cannot reduce this to $n^{2-\delta}$ for some absolute $\delta > 0$. Gowers and Long [55] conjectured that for sparser configurations, such a bound is attainable.

Conjecture 2.5. *For every $e \geq 3$, there is a $\delta = \delta(e) > 0$ so that $f_3(n, e + 4, e) = O(n^{2-\delta})$.*

In the previous subsection we discussed the approximate versions of BESC obtained in [26] and [87]. The following result of [41] gives an analogous approximate version of the Gowers–Long conjecture.

Theorem 2.6. *For every $e \geq 3$, there is a $\delta = \delta(e) > 0$ so that*

$$f_3\big(n, e + O(\log e), e\big) = O(n^{2-\delta}).$$

Recall that prior to Theorem 2.1 of [26], the state-of-the-art result for the BESC was inequality (2.1) of Sárközy–Selkow [87]. The above theorem then shows that with an error term close to that of Sárközy–Selkow, one can in fact improve the $o(n^2)$ bound on the number of edges to $O(n^{2-\delta})$.

As we noted after (2.2), one of the surprising implications of the (6, 3)-theorem is Roth's theorem. A result of this nature due to Gowers and Long [55] then states that a positive answer to Conjecture 2.5 for $e = 5$ would imply that for some $c > 0$, every $S \subseteq [n]$ of size n^{1-c} contains a nontrivial solution (i.e., one where not all integers are equal) of the equation $2x_1 + 2x_3 = x_3 + 3x_4$. This is related to a famous problem of Ruzsa [84], asking if for every linear equation E of the form $\sum_i a_i x_i = 0$ with $\sum_i a_i = 0$, and such that no (nonempty) proper subset of the coefficients a_i sums to 0, there is $X \subseteq [n]$ of size $n^{1-o(1)}$ without a nontrivial solution of E. As of now, a positive answer is only known when all but one of the a_i's are positive, via a straightforward generalization of Behrend's construction [16]. It would be very interesting to show that the relation between Ruzsa's problem and Conjecture 2.5 can be extended to other equations.

2.4. A Ramsey variant of BESC

Given the difficulty of the BESC, it is natural to look at various relaxations of it. For example, instead of looking at arbitrary r-graphs, one can look at those arising from a group (see [74] and its references). We now state another natural simplification of BESC that was recently suggested by Conlon and Nenadov. We say that a linear r-graph is *complete* if every pair of vertices of V belongs to exactly one edge of E. Such an object is sometimes called an r-Steiner system (when $r = 3$ this is a *Steiner triple system*). Conlon and Nenadov then suggested the following weaker Ramsey-type version of the BESC, namely proving that the following holds for every $r \geq 3$, $e \geq 3$, $c \geq 2$, and large enough $n \geq n_0(r, e, c)$: If $\mathcal{H}$ is an n-vertex complete linear r-graph then in every c-coloring of its edges one can find e edges of the same color, which are spanned by at most $(r-2)e + 3$ vertices. Note that BESC implies the above statement, as it gives the required monochromatic configuration in the most popular color. It is easy to see that this problem has a positive answer when $c = 1$ or when $e = 3$. The problem is wide open already when $e = 4$. The following result of [90] gives a positive answer to the Conlon–Nenadov problem assuming r is large enough.

Theorem 2.7. *For every integer c, there exists an $r_0 = r_0(c)$ such that for every $r \geq r_0$ and integer $e \geq 3$ there exists $n_0 = n_0(c, r, e)$ such that every c-coloring of a complete linear r-graph on $n > n_0$ vertices contains a monochromatic $((r-2)e + 3, e)$-configuration.*

In the natural case $c = 2$, it was further shown in [90] that $r_0(2) \leq 4$. The above results were recently generalized in [64], building on some of the tools and ideas introduced in [90]. One of these tools was an auxiliary graph which helps one "grow" $((r-2)e + 3, e)$-configurations, for $e = 3, 4, \ldots$, and thus prove Theorem 2.7. Perhaps one take-home message of [90] is that even when considering the Ramsey relaxation of the BESC, and even after adding the assumption that $r \geq r_0(c)$, one still has to work quite hard in order to find the $((r-2)e+3, e)$-configurations of the BESC.

3. VARIANTS OF THE REGULARITY LEMMA AND THEIR APPLICATIONS

In this section we discuss several variants of the regularity lemma and their relation to three of the most well-studied variants of the removal lemma. We will need the following definitions. For a fixed graph property $\mathcal{P}$, the *distance* of an n-vertex graph G from satisfying $\mathcal{P}$ is the smallest number of edges modifications (i.e., addition and removal) needed to turn G into an n-vertex graph satisfying $\mathcal{P}$. We say that an n-vertex graph is ε-*far* from satisfying $\mathcal{P}$ if G's distance from $\mathcal{P}$ is at least εn^2.

3.1. Triangle removal using weak regularity lemmas

In this subsection we focus on the triangle removal lemma, and more concretely, on the quantitative bounds for the function $\mathrm{Rem}(\varepsilon)$ introduced in Lemma 1.2. Actually, all the results will hold for the more general *graph removal lemma*, which states that for every graph H there is a function $\mathrm{Rem}_H(\varepsilon)$ so that if G is ε-far from being H-free then G contains $n^h/\mathrm{Rem}_H(\varepsilon)$ copies of H, where $h = |V(H)|$. In what follows, we will use $\mathrm{twr}(x)$ to denote the tower function, namely a tower of exponents of height x. For example, $\mathrm{twr}(3) = 2^{2^2}$.

The original proof of the triangle removal lemma in [85], and of its generalization to every fixed H [3], relied on Szemerédi's regularity lemma [95] and gave the bound $\mathrm{Rem}_H(\varepsilon) \leq M(\mathrm{poly}(\varepsilon))$. Let us sketch this proof when H is the triangle (the proof for general H is almost identical). Given G that is ε-far from being H-free, one first invokes the regularity lemma (with $\varepsilon/10$) in order to obtain an ε-regular partition of $V(G)$. One then removes from G edges that are either (i) inside one of the sets V_i, or (ii) connect V_i to V_j so that (V_i, V_j) is not $\varepsilon/10$-regular, or (iii) connect V_i to V_j so that $d(V_i, V_j) \leq \varepsilon/5$. Since this "cleaning process" removes less than εn^2 edges, at least one triangle remains in the new graph. By the nature of the cleaning process, there must be V_i, V_j, V_k so that this triangle has one vertex in each of these sets, and so that all three pairs of sets have density at least $\varepsilon/5$ and are $\varepsilon/10$-regular. One then invokes the so called *counting lemma* (see, e.g., Lemma 3.2 in [4]) in order to show that such a triple of sets V_i, V_j, V_k in fact contains $\mathrm{poly}(\varepsilon)|V_i||V_j||V_k|$ many triangles. Since each of these three sets has size at least $n/M(\varepsilon/10)$, we conclude that G has at least $n^3/M^3(\mathrm{poly}(\varepsilon))$ triangles. Unfortunately, Szemerédi's proof of his lemma gave the bound $M(\varepsilon) \leq \mathrm{twr}(\mathrm{poly}(1/\varepsilon))$, which combined with the preceding proof gives

$$\mathrm{Rem}_H(\varepsilon) \leq \mathrm{twr}\big(\mathrm{poly}(1/\varepsilon)\big). \tag{3.1}$$

As to lower bounds, extending a construction of [85] for triangle-freeness, Alon [1] proved that for every nonbipartite H, there is $C = C(H)$ satisfying

$$\mathrm{Rem}_H(\varepsilon) \geq (1/\varepsilon)^{C \log(1/\varepsilon)}. \tag{3.2}$$

There are two natural approaches for improving (3.1). The first would be to obtain better bounds for the regularity lemma, namely for $M(\varepsilon)$, which would immediately lead to improved bounds for $\operatorname{Rem}_H(\varepsilon)$. This, as well as numerous other applications of the regularity lemma, gave the hope that one can find a new proof of this lemma with significantly better quantitative bounds. These hopes were unfortunately shattered when Gowers [52] famously proved that $M(\varepsilon) \geq \operatorname{twr}(\operatorname{poly}(1/\varepsilon))$. The current record lower bound was obtained in [34] who showed that $M(\varepsilon) \geq \operatorname{twr}(1/\varepsilon^2)$, while a much shorter and simpler proof of Gowers's lower bound appears in [70].

Given the above, the second approach for improving (3.1) was to find a proof of the removal lemma that avoids the regularity lemma. This problem was open for many years until the breakthrough result of Fox [31], who found a new proof showing that

$$\operatorname{Rem}_H(\varepsilon) \leq \operatorname{twr}\big(O\big(\log(1/\varepsilon)\big)\big). \tag{3.3}$$

Fox's proof used an ad-hoc argument which was simplified in [25]. Given the simplicity of the proof of the removal lemma using the regularity lemma (sketched above), it is natural to ask if there is a weaker version of the regularity lemma, which is strong enough for proving the removal lemma (in a way similar to the original proof), yet weak enough so as to yield better bounds. Before describing two such proofs we should point out that the idea of devising weak regularity lemmas for specific applications was used before. Two notable such examples are the Frieze–Kannan *weak regularity lemma* [37] and the Duke–Lefman–Rödl [27] *cylinder lemma*. Below we describe two weaker versions of the regularity lemma which do produce bounds better than (3.1) by giving alternative proofs of (3.3).

Finding a regular partition of only part of the graph. Recall that when deriving the removal lemma from the regularity lemma, we only needed the 3 pairs among V_i, V_j, V_k to be ε-regular. The reason why the bound was so weak is that these sets are of size $n/\operatorname{twr}(\operatorname{poly}(1/\varepsilon))$. A special case of the cylinder lemma [27], shows that given an n-vertex graph, one can find three sets V_i, V_j, V_k so that each of the three pairs among them is ε-regular and the three sets have size $n/2^{\operatorname{poly}(1/\varepsilon)}$. Unfortunately, it does not appear that this lemma can be used to prove the triangle removal lemma since there is no structural connection between the 3 sets and G. In their work on the induced removal lemma (see Section 3.2), Alon, Fischer, Krivelevich, and Szegedy [4] proved the following related theorem.

Theorem 3.1. *For every $\varepsilon > 0$ and h, there is an $S = S(\varepsilon, h)$ so that every graph G has an equipartition $\{V_1, \ldots, V_k\}$ and a collection of subsets $W_1 \subseteq V_1, \ldots, W_k \subseteq V_k$ satisfying:*

(i) $\sum_{1 \leq i < j \leq k} |d(V_i, V_j) - d(W_i, W_j)| \leq \varepsilon k^2$;

(ii) *All pairs (W_i, W_j) are ε^h-regular;*

(iii) $|W_i| \geq |V(G)|/S$.

Furthermore, we have $\operatorname{Rem}_H(\varepsilon) \leq S(\operatorname{poly}(\varepsilon), h)$, where $h = |V(H)|$.

We note that the proof of the "furthermore" part of this theorem is almost identical to the way one derives the removal lemma from the regularity lemma, as we sketched above.

Note that item (i) gives us the required relation between G and the sets $W_1, \dots, W_k$, which is missing when applying the cylinder lemma [27].

Alon et al. [4] obtained a wowzer-type upper bound for $S(\varepsilon, h)$, where wowzer is the iterated version of the tower function. This wowzer-type bound resulted from using the *strong regularity lemma* which was introduced in [4]. It was later proved [24, 62] that the wowzer-type bounds for the strong regularity lemma are unavoidable, thus ruling out the possibility of improving the bounds for Theorem 3.1 by improving the bounds for the strong regularity lemma. A new proof of Theorem 3.1 was obtained by Conlon and Fox [24] who used the cylinder lemma [27] in a sophisticated way in order to prove that $S(\varepsilon, h) \leq \mathrm{twr}(\mathrm{poly}(1/\varepsilon))$. Note that even this improved bound does not give an improvement over (3.1), and as we mention below in Theorem 3.6, for *general graphs* this improved bound is the best possible. Hence, it appears as if Theorem 3.1 cannot be used to improve (3.1). However, by combining the ideas of [24] with those of [71], the following result was obtained in [86].

Theorem 3.2. *If G in Theorem 3.1 has $O(\varepsilon n^2)$ edges then $S(\varepsilon, h) \leq \mathrm{twr}(O(\log(1/\varepsilon)))$.*

Using the theorem above, and a minor variant of the proof of the removal lemma from the regularity lemma sketched above, one obtains (3.3) but only for graphs with $O(\varepsilon n^2)$ edges. So to reprove (3.3) in full generality it remains to prove that if G is ε-far from being H-free, then G has a subgraph G' with $O(\varepsilon n^2)$ edges which is $\Omega(\varepsilon)$-far from being H-free. Indeed, we could then apply the statement that holds only for graphs with $O(\varepsilon n^2)$ edges, and then use the fact that every copy of H in G' is also a copy of H in G. To find such a G', we first note that if G has δn^2 edge-disjoint copies of H, then G is clearly δ-far from being H-free, and that conversely, if G is ε-far from being H-free, then it contains at least $\varepsilon n^2/h^2$ edge-disjoint copies of H (where $h = |V(H)|$). Hence, taking G' to be the union of these edge-disjoint copies of H we obtain the required subgraph of G.

Modifying the graph. Recall that when proving the removal lemma, we first obtained an ε-regular partition of the graph, then removed edges from G, and then found many triangles in the new graph G'. Since we are already finding triangles in a modified version of G, one can ask if instead of finding a regular partition of G (which might be hard by Gowers's lower bound), it is enough to find a regular partition of a modified version of G. A version of the regularity lemma called the *regular approximation lemma* achieves this task.

Theorem 3.3. *For every $\varepsilon, \delta > 0$, there is a $T = T(\varepsilon, \delta)$ so that one can add/remove from an n-vertex graph at most δn^2 edges so that the new graph G' has a partition of order at most T in which* all *pairs are ε-regular.*

The first proofs of this lemma [69, 78] supplied (at best) only wowzer-type bounds. A better tower-type bound was obtained by Conlon and Fox [24]. Interestingly, the tower dependence is only on δ and not on ε. Unfortunately, for proving the removal lemma, one has to take $\delta \approx \varepsilon$, and so we again obtain (3.1). However, the following result of [71], which is a variant tailored for sparse graphs and appropriately dubbed the *sparse regular approximation lemma*, supplies a better bound.

Theorem 3.4.

Suppose G is a graph with $O(\varepsilon n^2)$ edges. Then one can add/delete from G at most $\varepsilon n^2/100$ edges so that the resulting graph G' has an ε^3-regular partition of order at most $\mathrm{twr}(O(\log 1/\varepsilon))$.

Let us briefly describe how the above theorem can be used to prove (3.3). First, as described after Theorem 3.2, it is enough to consider only graphs with $O(\varepsilon n^2)$ edges. Given such a G, we apply Theorem 3.4 to obtain G'. Since G' was obtained using few edge modifications, it is $\varepsilon/2$-far from being triangle free. We can now repeat the same argument used to derive the removal lemma from the regularity lemma, to infer that G' contains $n^h/\mathrm{twr}(O(\log 1/\varepsilon))$ copies of H (the improved bound comes from Theorem 3.4). Since we are allowed to add edges to G when producing G', one needs to be careful here since G' might contain "ghost" triangles that do not belong to G. However, it is not hard to show that at least half of the triangles in G' also belong to G thus completing the proof.

3.2. Improved bounds for the induced removal lemma?

We now consider the so called *induced removal lemma*, which is the induced variant of the removal lemma we discussed above. It states that for every fixed graph H, there is $\mathrm{Rem}^*_H(\varepsilon)$ so that if G is ε-far from being induced H-free then G has at least $n^h/\mathrm{Rem}^*_H(\varepsilon)$ induced copies of H. The fact that for every H such a function $\mathrm{Rem}^*_H(\varepsilon)$ exists was first obtained in [4] using Theorem 3.1. More precisely, what they proved was that $\mathrm{Rem}^*_H(\varepsilon) \le S(\mathrm{poly}(\varepsilon), h)$. As we noted in the previous subsection, a tower-type bound for Theorem 3.1 was obtained in [24] giving the improved $\mathrm{twr}(\mathrm{poly}(1/\varepsilon))$ upper bound for the induced removal lemma. Conlon and Fox later raised [25] the following natural problem, asking if one can extend (3.3) to the more difficult setting of the induced removal lemma.

Problem 3.5. Show that for every H we have $\mathrm{Rem}^*_H(\varepsilon) \le \mathrm{twr}(O(\log(1/\varepsilon)))$.

Since we know that $\mathrm{Rem}^*_H(\varepsilon) \le S(\mathrm{poly}(\varepsilon), h)$, it is natural to try and resolve the above problem by further reducing the upper bound for Theorem 3.1 to $\mathrm{twr}(O(\log(1/\varepsilon)))$. Recall that Theorem 3.2 shows that such a bound *is* attainable for graphs with $O(\varepsilon n^2)$ edges, implying a positive answer for Problem 3.5 for graphs with this many edges. Unfortunately, a recent result of [67] shows that such a bound is not attainable in general.

Theorem 3.6. *There is a* $\mathrm{twr}(\mathrm{poly}(1/\varepsilon))$ *lower bound for* $S(\varepsilon, 10)$ *in Theorem* 3.1.

Another approach for resolving Problem 3.5 is to reduce the general case of bounding $\mathrm{Rem}^*_H(\varepsilon)$ to the case where G has $O(\varepsilon n^2)$ edges, since for this special case we can resolve Problem 3.5. As we observed after Theorem 3.2, such a reduction is easy to obtain for the (noninduced) removal lemma. There is a very natural way to try and extend that argument to the setting of induced H-freeness. We say that two induced copies of H in G are *pair-disjoint* if they share at most one vertex. As in the case of H-freeness, it is clear that if G contains εn^2 pair-disjoint induced copies of H then G is ε-far from being induced H-free. Perhaps surprisingly, the converse is not true. For example, one can construct a graph that is

ε-far from being induced C_4-free, yet it contains only $O(\varepsilon^2 n^2)$ pair-disjoint induced copies of C_4. However, it is natural to ask if the following approximate result holds.

Problem 3.7. Show that if G is ε-far from being induced H-free then G contains at least $\mathrm{poly}(\varepsilon) \cdot n^2$ pair-disjoint induced copies of H.

It is a simple corollary of the induced removal lemma itself that if G is ε-far from being induced H-free then G contains $\Omega(n^2)$ pair-disjoint copies of H, but the hidden constant has a tower-type dependence on ε. The question is if this can be made polynomial in ε. Besides being a natural problem, solving Problem 3.7 would also lead to a solution of Problem 3.5. This can be proved using the ideas of [71]. A much simpler argument was noted independently by Jacob Fox (private communication).

3.3. The hypergraph regularity lemma

The following lemma is the natural generalization of Lemma 1.2 (the triangle removal lemma) to r-graphs. Here, $K_{r+1}^{(r)}$ denotes the complete r-graph on $r+1$ vertices, namely, a set of $r+1$ vertices containing all possible $r+1$ r-edges (so $K_3^{(2)}$ is a triangle).

Theorem 3.8 (Hypergraph removal lemma). *For every $r \geq 2$ and $\varepsilon > 0$, there is an $\mathrm{Rem}_r(\varepsilon)$ so that the following holds. Suppose $\mathcal{H}$ is an n-vertex r-graph with the property that one should remove at least εn^r of its edges to make it $K_{r+1}^{(r)}$-free. Then $\mathcal{H}$ contains at least $n^{r+1}/\mathrm{Rem}_r(\varepsilon)$ copies of $K_{r+1}^{(r)}$.*

The first to conjecture the above extension of the triangle removal lemma were Erdős, Frankl, and Rödl [29] in the 1980s. One of the main motivations for obtaining Theorem 3.8 was the observation of Frankl and Rödl [36] (see also [92]) that it would give an alternative proof of Szemerédi's theorem for progressions of arbitrary length. Another motivation was the hope that it would lead to a solution of Conjecture 1.1.

As we discussed earlier, the proof of the triangle removal lemma relied on Szemerédi's regularity lemma. The quest for an r-graph regularity lemma that would allow one to prove Theorem 3.8 took about 20 years. The first milestone was the result of Frankl and Rödl [36], who obtained a regularity lemma for 3-graphs and using it proved Theorem 3.8 for $r = 3$. About 10 years later, the approach of [36] was extended to r-graphs (for arbitrary $r \geq 2$) by Rödl, Skokan, Nagle, and Schacht [73, 81]. At the same time, Gowers [54] obtained an alternative version of the regularity lemma for r-graphs. Shortly after, Tao [97] and Rödl and Schacht [77,78] obtained two additional versions of the lemma. A more detailed discussion appears in [75].

For the next discussion, we need to introduce the functions comprising the Ackermann hierarchy. Set $A_1(x) = 2^x$ and, for every $r \geq 2$, define the function $A_r(x)$ to be the result of iterating A_{r-1} on itself x times. So $A_2(x)$ is the tower function and $A_3(x)$ is the wowzer function. We refer to A_r as the rth Ackermann function.

Although the above mentioned regularity lemmas for r-graphs are quite different from each other, they all involve constants that grow as fast as the rth Ackermann function.

As a result, all proofs of the removal lemma for r-graphs give (at best) a bound of the form $\mathrm{Rem}_r(\varepsilon) \le A_r(1/\varepsilon)$. This leads to the following open problem:

Problem 3.9. Obtain primitive recursive bounds for the r-graph removal lemma. That is, show that there is a universal constant r_0 so that $\mathrm{Rem}_r(\varepsilon) \le A_{r_0}(1/\varepsilon)$ for every $r \ge 2$.

It is natural to first ask if one can resolve the above problem simply by improving the constants involved in one of the r-graph regularity lemmas mentioned above, which are known to imply the removal lemma. As mentioned above, Gowers [52] proved that for $r = 2$, one cannot obtain better than $A_2(\mathrm{poly}(1/\varepsilon))$ bounds for the graph regularity lemma. This result was extended to all $r \ge 2$ in [72].

Theorem 3.10. *For every $r \ge 2$, there is an $A_r(\log(1/\varepsilon))$ lower bound for the r-graph regularity lemma.*

Hence, if one wishes to improve the A_r-type bounds for the r-graph removal lemma, one has to develop new variants of the r-graph regularity lemma that, on the one hand, are strong enough to prove Theorem 3.8, and, on the other hand, are weak enough to yield better bounds.

The main challenge in proving Theorem 3.10 is facilitating an inductive approach (on r): one has to prove a stronger lower bound, showing that even very weak versions of the r-graph regularity lemma cannot give bounds better than $A_r(\log(1/\varepsilon))$. Among other things, one has to show that the lower bound holds even if one is allowed to change, say, a 0.01-fraction of the edges. As the reader might recall, this is exactly the type of regularity lemma mentioned in Theorem 3.4, where we stated an upper bound for such a weak version of the lemma. As part of the proof in [72], the following matching lower bound was obtained.

Theorem 3.11. *The $\mathrm{twr}(O(\log 1/\varepsilon))$ upper bound in Theorem 3.4 is tight.*

Returning to the discussion in Section 3.1, Theorem 3.11 implies that the approach described prior to Theorem 3.4 cannot improve (3.3).

4. VARIANTS OF THE REMOVAL LEMMA

We next describe several problems related to the triangle removal lemma mentioned in Section 1. Since its inception [85], the triangle removal lemma was extended in various ways. Two such generalizations, the general removal lemma and the induced removal lemma, were discussed in the previous section. These extensions culminated in the following result of [9], where a graph property is *hereditary* if it is closed under removal of vertices.

Theorem 4.1. *For every hereditary property $\mathcal{P}$, there is a function $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$, so that if a graph G is ε-far from satisfying $\mathcal{P}$, then a random and uniform sample of $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$ vertices from G spans a graph not satisfying $\mathcal{P}$ with probability at least $2/3$.*

Given a (possibly infinite) family of graphs $\mathcal{F}$, let $\mathcal{P}^*_{\mathcal{F}}$ denote the property of being induced $\mathcal{F}$-free, namely, not containing an induced copy of any $F \in \mathcal{F}$. It is clear that the family of hereditary properties coincides with the family of properties $\mathcal{P}^*_{\mathcal{F}}$, so Theorem 4.1 is the most general version of the removal lemma one can hope to prove. The fact that Theorem 4.1 is indeed a generalization of the removal lemma and the induced removal lemma follows from the reasoning in the paragraph following the statement of Theorem 1.2. The reason we change gears here is that, when $\mathcal{F}$ is infinite, stating the removal lemma for $\mathcal{P}^*_{\mathcal{F}}$ in the style of Theorem 1.2 becomes cumbersome. The same applies to Problem 4.2 below.

4.1. A theoretical computer science interlude

Although the removal lemmas we discuss below have purely combinatorial statements, part of the motivation leading to these results came from questions in theoretical computer science, more specifically from the area of *graph property testing* [48]. The interplay between this area and extremal combinatorics has been extremely fruitful, with many questions raised in one area motivating the development of new tools in the other. Examples of tools are the weak regularity lemma of Frieze and Kannan [37], the conditional regularity lemma of Alon, Fischer, and Newman [5], and the notion of partition oracles [59]. A comprehensive discussion on the combinatorial aspects can be found in Lovász's book [68] and on the more algorithmic aspects in Goldreich's book [48]. In this subsection we give a brief background on this area.

Classical models of computation ask for an algorithm that can decide if an input satisfies some property $\mathcal{P}$, for example, whether an input graph G is planar, or whether an input matrix A is invertible. It is easy to see that in this case we have to read the entire input at least once, for example, because deleting a single edge of the graph might make it planar, or because changing a single entry of A might make it invertible. Due to the need to analyze huge inputs, which might be too costly to scan even once, researchers introduced a new type of algorithms, called *property testers*, that solve only relaxed versions of the classical decision problem, but do so extremely fast. These are randomized algorithms whose goal is to distinguish (with high probability, say, $2/3$) between objects satisfying some fixed property $\mathcal{P}$ and those that are ε-far from satisfying it. The study of such problems originated in the seminal papers of Rubinfeld and Sudan [83], Blum, Luby, and Rubinfeld [18], and Goldreich, Goldwasser, and Ron [50]. Below are the precise definitions related to property testing of graphs.

We say that a graph property is *testable* if there is a function $q_{\mathcal{P}}(\varepsilon)$ so that by sampling a set of vertices S of size $q_{\mathcal{P}}(\varepsilon)$ from a graph G, one can distinguish with probability at least $2/3$ between the following two cases (i) G satisfies $\mathcal{P}$ and (ii) G is ε-far from $\mathcal{P}$. So the fact that $\mathcal{P}$ is testable means that we can distinguish a graph $G \in \mathcal{P}$ from G that is ε-far from $\mathcal{P}$ while looking at a subgraph of G of constant size! It is not hard to see that the triangle removal lemma is equivalent to the statement that triangle-freeness is a testable property, and that bounding $\mathrm{Rem}(\varepsilon)$ is equivalent to bounding the corresponding function $q(\varepsilon)$.

Some of the most important questions in property testing are those asking if general families of properties are testable. Observe that Theorem 4.1 implies that every hereditary graph property $\mathcal{P}$ is testable. Indeed, the algorithm for testing $\mathcal{P}$ simply samples a set S of $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$ vertices. If the graph spanned by S satisfies $\mathcal{P}$ then the algorithm declares that the input satisfies $\mathcal{P}$, otherwise it declares the input is ε-far from $\mathcal{P}$. Since $\mathcal{P}$ is hereditary, if G satisfies $\mathcal{P}$, the algorithm will declare this with probability 1. On the other hand, the definition of $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$ guarantees that if G is ε-far from $\mathcal{P}$, the algorithm will declare this with probability at least $2/3$.

As we noted above, the algorithm for testing a hereditary property always answers correctly when the input belongs to $\mathcal{P}$. Such an algorithm is said to have one-sided error. It was shown in [9] that hereditary properties are (essentially) the only properties that can be tested by a one-sided error algorithm. A characterization of the properties that can be tested in the more general setting of two-sided error algorithms was obtained in [6] and [20]. These results were extended to r-graphs in [15,61,79].

It should be noted that while the algorithms defined above have running time that depends only on ε (and are independent of $|V(G)|$), the dependence on ε might be enormous. Indeed, as we discussed in Section 3.1, even in the special case of $\mathcal{P}$ being triangle-freeness, the running time of the testing algorithm is given by the tower-type function in (3.3). Furthermore, a result of [10] shows that there are properties $\mathcal{P}$ for which $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$ grows faster than any recursive function.

While the results discussed above give rather satisfactory *qualitative* answers, by the previous paragraph they give very poor *quantitative* answers. Hence, once we know that a property is testable, the next natural question is whether we can obtain a "reasonable" bound for $q_{\mathcal{P}}(\varepsilon)$. As in many questions, the natural definition of reasonable is polynomial. We thus say that $\mathcal{P}$ is *easily testable* if it is testable with a polynomial sample, that is, if $q_{\mathcal{P}}(\varepsilon) = \mathrm{poly}(1/\varepsilon)$. One of the most important open problems in this area was popularized by Goldreich [48] and by Alon and Fox [7], who asked for a characterization of the easily testable graph properties. Currently, this problem is open even when restricted to hereditary properties. Note that, by the above discussion, proving that a hereditary property $\mathcal{P}$ is easily testable is equivalent to proving that in Theorem 4.1 we have $\mathrm{Rem}_{\mathcal{P}}(\varepsilon) = \mathrm{poly}(1/\varepsilon)$. This leads to the following open problem.

Problem 4.2. Characterize hereditary graph properties $\mathcal{P}$ for which $\mathrm{Rem}_{\mathcal{P}}(\varepsilon) = \mathrm{poly}(1/\varepsilon)$.

This line of research was initiated by Alon [1], who proved that if $\mathcal{P}_H$ is the property of being H-free, then $\mathcal{P}_H$ is easily testable if and only if H is bipartite. Another notable early result was obtained by Goldreich, Goldwasser, and Ron [50] who proved that for any fixed k, the property of being k-colorable is easily testable. This was a major improvement over an earlier result of Rödl and Duke [76] who used the regularity lemma and (implicitly) gave a tower-type upper bound for testing k-colorability. In the next subsection we discuss recent progress related to Problem 4.2.

We finally mention another variant of Theorem 4.1. It is natural to ask if using a sample of vertices of constant size, one can not only detect if an input G is ε-far from $\mathcal{P}$, but further *estimate* G's distance from $\mathcal{P}$. In the literature on property testing, this is called *tolerant testing*. Such a result was obtained for monotone properties in [11], and for all testable properties, and in particular all hereditary properties, in [30]. A recent result of [60] further shows how to efficiently transform any bound for $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$ into a bound for tolerantly testing $\mathcal{P}$.

4.2. Removal lemmas with polynomial bounds

In this subsection we describe the progress towards Problem 4.2. It will be more convenient to think of a hereditary property in terms of its forbidden induced subgraphs, that is, represent it as $\mathcal{P}^*_{\mathcal{F}}$, as defined after Theorem 4.1. When $\mathcal{F}$ consists of a single graph F we will use the notation $\mathcal{P}^*_F$.

We first consider hereditary properties $\mathcal{P}^*_{\mathcal{F}}$ with finite $\mathcal{F}$. Recall that a graph F is *bipartite* if $V(F)$ can be partitioned into two sets A, B that are both independent, that is, contain no edges. A graph F is *cobipartite* if $V(F)$ can be partitioned into two complete graphs A, B. Finally, F is *split* if $V(F)$ can be partitioned into A, B, one independent and the other complete. The following result is proved in [45].

Theorem 4.3. *If $\mathcal{F}$ is a finite family of graphs that contains a bipartite graph, a cobipartite graph and a split graph then $\mathcal{P}^*_{\mathcal{F}}$ is easily testable.*

As discussed in [45], many known and new results can be derived from Theorem 4.3. For example, Alon and Fox [7], using a somewhat involved ad hoc argument, proved that the property of being induced P_4-free (P_4 is the path on 4 vertices) is easily testable. This follows immediately from Theorem 4.3 since P_4 is bipartite, cobipartite, and split.

The next theorem from [45] shows that the sufficient condition in Theorem 4.3 is almost necessary.

Theorem 4.4. *Let $\mathcal{F}$ be a finite family for which $\mathcal{P}^*_{\mathcal{F}}$ is easily testable. Then $\mathcal{F}$ contains a bipartite graph and a cobipartite graph.*

As in the case of Theorem 4.3, the above theorem can also be used in order to obtain many previous results showing that certain properties are not easily testable. Having given both a necessary and a sufficient condition, it is natural to ask if one of them in fact characterizes the finite families $\mathcal{F}$ for which $\mathcal{P}^*_{\mathcal{F}}$ is easily testable. Unfortunately, it was proved in [45] that none of them is a characterization. Hence, even the special case of Problem 4.2, that of characterizing the finite families of graph $\mathcal{F}$ for which $\mathcal{P}^*_{\mathcal{F}}$ is easily testable, is still open.

In addition to the above results concerning finite $\mathcal{F}$, [45] also obtained a sufficient condition guaranteeing that $\mathcal{P}^*_{\mathcal{F}}$ is easily testable for general families $\mathcal{F}$. Instead of describing this condition, we discuss a corollary of it, which concerns the family of *semialgebraic* graph properties. A semialgebraic graph property $\mathcal{P}$ is given by an integer $k \geq 1$, a set of real $2k$-variate polynomials $f_1, \dots, f_t \in \mathbb{R}[x_1, \dots, x_{2k}]$ and a Boolean function $\Phi : \{\text{true}, \text{false}\}^t \to \{\text{true}, \text{false}\}$. A graph G satisfies a property $\mathcal{P}$ if one can assign a point $p_v \in \mathbb{R}^k$ to each vertex $v \in V(G)$ in such a way that a pair of distinct vertices u, v are adjacent if and only if

$$\Phi\big(f_1(p_u, p_v) \geq 0, \dots, f_t(p_u, p_v) \geq 0\big) = \text{true}.$$

In the expression $f_i(p_u, p_v)$, we substitute p_u into the first k variables of f_i and p_v into the last k variables of f_i.

Some examples of semialgebraic graph properties are those that correspond to being an intersection graph of certain semialgebraic sets in $\mathbb{R}^k$. For example, a graph is an *interval graph* if one can assign an interval in $\mathbb{R}$ to each vertex so that u, v are adjacent iff their intervals intersect. Similarly, a graph is a *unit disc graph* if it is the intersection graph of unit discs in $\mathbb{R}^2$.

The family of semialgebraic graph properties has been extensively studied by many researchers, see, e.g., [35] and its references. Alon conjectured that every semialgebraic graph property is easily testable. This conjecture was verified in [45].

Theorem 4.5. *Every semialgebraic graph property is easily testable.*

The proofs of Theorems 4.3 and 4.5 use the *conditional regularity lemma* of Alon, Fischer, and Newman [5]. This variant of the regularity lemma states that if there is a fixed bipartite graph H so that the graph G has no induced copy of H (when considering only the edges connecting the two sided of H), then G has an ε-regular partition of size only $\text{poly}(1/\varepsilon)$. In fact, the same statement holds if G has only a few copies of H. One of the key steps in the proofs of Theorems 4.3 and 4.5 is then to show that for every relevant property $\mathcal{P}$, an appropriate H as above exists. A related strategy was taken in [32] in the setting of tournaments.

We conclude this subsection with a problem of Alon [1], who asked to characterize the graphs H for which the property of being induced H-free is easily testable. It can be easily checked that Theorems 4.3 and 4.4 can be used to answer this question for all graphs except C_4. There is an interesting reason why this case remains elusive. As we mentioned in the previous paragraph, in the proof of Theorem 4.3 we use the fact that graphs satisfying the properties in its statement are guaranteed to have ε-regular partitions of order $\text{poly}(1/\varepsilon)$. The reason why induced C_4-freeness is harder is that a graph might satisfy this property and still only have regular partitions as in Gowers's example [52] (i.e., having tower-type size). To see this, one just has to note that every split graph (defined before Theorem 4.3) is induced C_4-free, and that one can assume that Gowers's example is a bipartite graph. Hence, taking this graph, and turning one of its independent sets into a complete graph, gives the required example.

Alon and Fox [7] asked if one can improve the tower-type bounds for induced C_4-freeness, which follow from the bound on $\text{Rem}^*_H(\varepsilon)$ discussed in Section 3.2. The following result of [42] improved this to a mere exponential bound.

Theorem 4.6. $\text{Rem}_{C_4}(\varepsilon) \leq 2^{\text{poly}(1/\varepsilon)}$.

The problem of improving this bound to $\text{poly}(1/\varepsilon)$ remains open.

4.3. Removal lemmas of prescribed growth and generalized Turán problems

In the previous parts of this paper, we have mentioned that there are various types of lower and upper bounds for the function $\text{Rem}_{\mathcal{P}}(\varepsilon)$. However, in all cases we could either prove that this function is polynomial (as in the previous subsection) or we had a huge tower-type difference between the best lower and upper bounds (e.g., when $\mathcal{P}$ is triangle-freeness, see (3.1) and (3.2)). This raises the natural question of finding, for a given growth function f, a property $\mathcal{P}$ for which $\mathrm{Rem}_{\mathcal{P}}(\varepsilon) \approx f(\varepsilon)$. A further motivation for this problem comes from theoretical computer science (see Section 4.1). One of the most basic results in this area is the *time hierarchy theorem*, stating (roughly) that for every (natural) function f, there are computational tasks requiring time $f(n)$ on inputs of size n. There are other theorems of this type with respect to memory usage, random bits, etc. Goldreich [48] asked for such a hierarchy theorem for the query complexity of testing graph properties. As we discussed in Section 4.1, the query complexity of testing a hereditary $\mathcal{P}$ with one-sided error is given by $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$. Hence, the following theorem from [43] gives a hierarchy theorem for testing graph properties with one-sided error.

Theorem 4.7. *For every decreasing $f : (0,1) \to \mathbb{N}$ satisfying $f(x) \geq 1/x$, there is a hereditary graph property $\mathcal{P}$ satisfying $f(\varepsilon) \leq \mathrm{Rem}_{\mathcal{P}}(\varepsilon) \leq \varepsilon^{-14} f(\varepsilon/c)$, where c is an absolute constant.*

As an immediate application of the above theorem, we see that there is a property $\mathcal{P}$ for which $\mathrm{Rem}_{\mathcal{P}}(\varepsilon) = 2^{\Theta(1/\varepsilon)}$ or one for which $\mathrm{Rem}_{\mathcal{P}}(\varepsilon) = \mathrm{twr}(\Theta(1/\varepsilon))$. The properties used in the proof of Theorem 4.7 are quite simple. Given f, the property $\mathcal{P}$ is that of not containing a cycle whose length belongs to the set of integers $\{a_1, a_2, \ldots\}$ where $a_1 = 3$ and for every $i \geq 1$ we define $a_{i+1} = 2f(1/2(a_i + 2)^2) + 1$.

While the properties used in the proof of Theorem 4.7 are simple, the proof that they satisfy its assertion is more complicated, and relies on a theorem we describe below. Turán's Theorem [100], one of the cornerstone results in graph theory, determines the maximum number of edges in an n-vertex graph that does not contain a K_t (the complete graph on t vertices). Turán's problem is the following more general question: for a fixed graph H and an integer n, what is the maximum number of edges in an n-vertex H-free graph? This quantity is denoted by $\mathrm{ex}(n, H)$. Estimating $\mathrm{ex}(n, H)$ for various graphs H is one of the most well-studied problems in graph theory. Alon and Shikhelman [12] have recently initiated the systematic study of the following natural generalization of $\mathrm{ex}(n, H)$; for fixed graphs H and T, estimate $\mathrm{ex}(n, T, H)$, which is the maximum number of copies of T in an n-vertex graph that contains no copy of H. Note that $\mathrm{ex}(n, H) = \mathrm{ex}(n, K_2, H)$. For the sake of brevity, we refer the reader to [12] for more background and motivation. Let us just mention that this family of problems is also related to those discussed in Section 2 since it is not hard to see that if we set D to be the graph comprising of two triangles sharing an edge, then $\mathrm{ex}(n, K_3, D) = \Theta(f_3(n, 6, 3))$.

Some of the most well-studied graphs analyzed in the setting of Turán problems are cycles. In the setting of generalized Turán problems, Bollobás and Györi [19] and Györi and Li [58] obtained tight bounds for $\mathrm{ex}(n, C_3, C_5)$ and $\mathrm{ex}(n, C_3, C_{2k+1})$. The main result of [43] was a tight bound for $\mathrm{ex}(n, C_k, C_\ell)$ for all pairs k, ℓ. For odd cycles, it states the following.

Theorem 4.8. *For every $2 \leq k < \ell$, we have $\mathrm{ex}(n, C_{2k+1}, C_{2\ell+1}) = \Theta_k(\ell^{k+1} n^k)$.*

4.4. Three generalizations of induced $\mathcal{F}$-freeness

Removal for linear combinations of subgraph statistics. For a fixed integer h, let us assign a weight $w_H \in [0,1]$ to every graph H on h vertices, and then collect all these weights into a sequence denoted $\overline{w}$. Let $d_H(G)$ denote the fraction of subsets of $V(G)$ of size h that induce a copy of H. Given h, a sequence of weights $\overline{w}$ as above, and $c \geq 0$, we say that a graph G satisfies $\mathcal{P}_{h,\overline{w},c}$ if $\sum_H w_H \cdot d_H(G) \leq c$. Note that if $c = 0$ and the only nonzero entry of $\overline{w}$ is w_H, then $\mathcal{P}_{h,\overline{w},c}$ is the property of being induced H-free. In a similar manner, for every finite family of graphs $\mathcal{F}$, we can encode the property of being induced $\mathcal{F}$-free. In particular, for every H, we can encode the property of being (not necessarily induced) H-free. Goldreich and Shinkar [51] conjectured that one can extend the induced removal lemma [4] by proving that every property $\mathcal{P}_{h,\overline{w},c}$ is testable. They in fact conjectured that these properties can be tested using a very restricted type of testing algorithm. A result of [47] shows that some of these properties are not testable at all.

Theorem 4.9. *There is a property $\mathcal{P}_{4,\overline{w},\frac{5}{16}}$ which is not testable with $n^{1/100}$ queries.*

To prove the above theorem, it is shown in [47] how to define a vector $\overline{w}$ so that the resulting property $\mathcal{P}_{4,\overline{w},\frac{5}{16}}$ encodes the property of being a quasirandom graph in the sense of Chung, Graham, and Wilson [23]. It remains an open problem to decide if the properties $\mathcal{P}_{h,\overline{w},c}$ can at least be tested using $o(n^2)$ edges queries.

Removal against an arbitrary distribution. Suppose G_1, G_2 are two graphs on the same vertex-set V and $\mathcal{D}$ is a distribution on V. The distance between G_1 and G_2 with respect to $\mathcal{D}$ is then defined to be $\sum_{\{x,y\} \in E(G_1) \Delta E(G_2)} \mathcal{D}(x) \cdot \mathcal{D}(y)$. We say that the pair $(G, \mathcal{D})$ is ε-far from satisfying a graph property $\mathcal{P}$ if for every $G' \in \mathcal{P}$, the distance between G and G' with respect to $\mathcal{D}$ is at least ε. Observe that the above definition generalizes the definitions we presented at the beginning of Section 3 which correspond to the uniform distribution over $V(G)$, that is, the one that assigns every vertex a weight of $1/n$. Now, for a given hereditary property $\mathcal{P}$ and $\varepsilon > 0$ we let $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ be the smallest integer so that for *every* distribution $\mathcal{D}$, if G is ε-far from $\mathcal{P}$ with respect to $\mathcal{D}$, then a sample of $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ vertices from $V(G)$, sampled *according to* $\mathcal{D}$, induces a graph not satisfying $\mathcal{P}$ with probability at least $2/3$. The order of quantifiers here is crucial; the definition of $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ requires that it would suffice for *every* $\mathcal{D}$. It is again clear that the above definition is much stronger than the one introduced at the beginning of Section 4, since $\mathrm{Rem}_{\mathcal{P}}(\varepsilon)$ only applies to the uniform distribution.

A priori it is not clear why a function $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ as above should exist for any (interesting) hereditary property. Goldreich [49] proved that such a function indeed exists for several types of hereditary properties. The main motivation for his study was that similar algorithmic tasks have been studied in many other settings, where they are called *distribution-free algorithms*, see [49] for more background. Goldreich asked if a function $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ as above exists for every hereditary $\mathcal{P}$. To answer this question, we need an important definition. We say that a graph property $\mathcal{P}$ is *extendable* if for every graph G satisfying $\mathcal{P}$, we can add to G a new vertex v and connect it to $V(G)$ in such a way that the resulting graph will also satisfy $\mathcal{P}$. The following answer to Goldreich's problem was given in [46].

Theorem 4.10.

If $\mathcal{P}$ is hereditary, then $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ *exists if and only if $\mathcal{P}$ is extendable.*

It was also proved in [46] that several natural restrictions on $\mathcal{D}$ guarantee that $\mathrm{Rem}'_{\mathcal{P}}(\varepsilon)$ exists for every hereditary $\mathcal{P}$. For example, this is the case if we assume that $\max_{v \in V(G)} \mathcal{D}(v) = o(1)$ or if we assume that $\min_{v \in V(G)} \mathcal{D}(v) = \Omega(1/|V(G)|)$. At a high level, the proof of Theorem 4.10 in [46] follows the framework of [9], but the fine details differ substantially. The proof in [9] uses Szemerédi's regularity lemma and its variants in order to handle every hereditary property, but only with respect to the uniform distribution. In [46] a new version of the regularity lemma is introduced, which takes into account the weight function $\mathcal{D}$, yet produces bounds that are independent of $\mathcal{D}$ (for extendable properties).

Removal for ordered graphs and matrices. For a fixed $k \times k$ matrix H with $0/1$ entries, we say that an $n \times n$ matrix A is H-free if there are no $r_1 < \cdots < r_k$ and $c_1 < \cdots < c_k$ so that $A_{r_i,c_j} = H_{i,j}$ for every $i, j \in [k]$. We define A to be ε-far from being H-free if one should change at least εn^2 of its entries to make it H-free. Observe that the matrix property of being H-free depends on the ordering of rows and columns. This is in sharp contrast to the graph property of being H-free which is independent of the "names" (or the ordering) of the vertices.

Alon, Fischer, and Newman [5] asked if the graph removal lemma can be extended to the setting of matrices, that is, if every A that is ε-far from being H-free contains at least $n^{2k}/\mathrm{Rem}_H(\varepsilon)$ copies of H. Alon, Ben-Eliezer, and Fischer [2] recently gave a positive answer to this question, showing that $\mathrm{Rem}_H(\varepsilon)$ can be bounded by a wowzer function of ε. Using the methods discussed in Section 3 this can probably be reduced to a tower-type bound. But the following problem is still open.

Problem 4.11. Obtain $\mathrm{poly}(1/\varepsilon)$ bounds for the matrix removal lemma.

We should point out that Alon, Fischer, and Newman [5] obtained a polynomial bound for the *unlabeled* variant of the matrix removal lemma, that is, one where the order of rows/columns of H does not matter. Equivalently, the result of [5] gives a polynomial bound for the induced removal lemma (see Section 3.2) in bipartite graphs. In this case the input G is an $n \times n$ bipartite graph, and G has an induced copy of a bipartite H on vertex sets U_1, U_2 if it has an induced copy in which $U_1 \subseteq V_1$ and $U_2 \subseteq V_2$, or vice versa. This efficient removal lemma was instrumental in the results described in Section 4.2.

4.5. Arithmetic removal lemmas for linear equations and functions

Considering the importance of the removal lemma, it is natural to ask if analogous results can be obtained in other settings. A notable example is the removal lemma for linear equations over groups obtained by Green [56]. A significantly simpler proof of Green's result was obtained by Král', Serra, and Vena [65] who derived it from the graph removal lemma. To state Green's result, we need a few natural generalizations of the notions we used in the setting of graphs. Let $S \subseteq [n]$ be a set of integers, let M be an $\ell \times t$ integer matrix, and $b \in \mathbb{N}^{\ell}$ an integer vector. We say that S is (M, b)-free if there is no $x \in S^t$ satisfying $Mx = b$, and say that S is ε-far from being (M, b)-free if we need to remove at least εn of its elements to make it (M, b)-free. Finally, we say that the pair (M, b) has the *removal property* if there is a function $\mathrm{Rem}_{M,b}(\varepsilon)$ so that if S is ε-far from being (M, b)-free, then S^t contains at least $n^{t-\ell}/\mathrm{Rem}_{M,b}(\varepsilon)$ vectors x satisfying $Mx = b$. Green's result then states that for $\ell = 1$ (i.e., for a single equation), every pair $(M, 0)$ has the removal property. He conjectured [56] that, for every M, the pair $(M, 0)$ has the removal property. Green's conjecture was verified in the following stronger form independently by [66] and [89]. Both proofs rely on Theorem 3.8.

Theorem 4.12. *Every pair (M, b) has the removal property.*

We conclude by describing an extension of Theorem 4.1 from the setting of graphs to the setting of boolean functions $f : \mathbb{F}_2^n \to \{0, 1\}$. Let $\mathcal{P}$ be a property of such functions, and say that f is ε-far from satisfying $\mathcal{P}$ if one should change the truth table of f in at least $\varepsilon 2^n$ places to make it satisfy $\mathcal{P}$. Let $\mathcal{T}$ be the property of such functions indicating that there is no pair $x, y \in \mathbb{F}_2^n$ so that $f(x) = f(y) = f(x + y) = 1$. Then Green's result [56] (mentioned above) implies that if f is ε-far from $\mathcal{T}$, then there are $2^{2n}/\mathrm{Rem}(\varepsilon)$ many pairs x, y witnessing this fact. Green's proof gave a tower-type bound for $\mathrm{Rem}(\varepsilon)$, which was improved to $\mathrm{poly}(1/\varepsilon)$ by [33], using tools related to the solution of the famous cap-set conjecture (see the discussion in [33]).

It is natural to ask for a unifying explanation for why property $\mathcal{T}$ above obeys a removal lemma. Such a systematic study was initiated by Kaufman and Sudan [63] who emphasized the role of invariance. Observe that a key feature in graph properties is that vertex names do not play a role, or more formally, they are closed under isomorphism. This motivated [17] to conjecture that a result analogous to Theorem 4.1 should hold in the setting of boolean functions. To state it we need two definitions. A property of boolean functions is *linear-invariant* if for every $f \in \mathcal{P}$ and any linear transformation $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ we have $f \circ L \in \mathcal{P}$ where $(f \circ L)(x) = f(L(x))$. We also say that a linear invariant $\mathcal{P}$ is *subspace-hereditary* if for every $f \in \mathcal{P}$ and every linear subspace U of $\mathbb{F}_2^n$ the restriction $f_{|U} \in \mathcal{P}$. We can thus think of linear-invariant subspace-hereditary properties as the analogue of hereditary graph properties. To further emphasize this analogy, it was observed in [17] that just as hereditary properties are those characterized by a (possibly infinite) family of forbidden induced subgraphs (as discussed after Theorem 4.1), then every linear-invariant subspace-hereditary property can be characterized by a (possibly infinite) family of forbidden "patterns" like the one forbidden in the above $\mathcal{T}$. The conjecture raised in [17] was that every linear-invariant subspace-hereditary property of boolean functions has a removal lemma. This conjecture was recently verified by Tidor and Zhao [99].

ACKNOWLEDGMENTS

I am grateful to Noga Alon, Lior Gishboliner, and Yevgeny Levanzov for helpful remarks on a preliminary version of this manuscript.

FUNDING

Supported in part by ISF Grant 1028/16, ERC Consolidator Grant 863438 and NSF-BSF Grant 2019.

REFERENCES

[1] N. Alon, Testing subgraphs in large graphs. *Random Structures Algorithms* **21** (2002), 359–370.

[2] N. Alon, O. Ben-Eliezer, and E. Fischer, Testing hereditary properties of ordered graphs and matrices. In *IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 848–858, 2017.

[3] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, The algorithmic aspects of the Regularity Lemma. *J. Algorithms* **16** (1994), 80–109.

[4] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, Efficient testing of large graphs. *Combinatorica* **20** (2000), 451–476.

[5] N. Alon, E. Fischer, and I. Newman, Testing of bipartite graph properties. *SIAM J. Comput.* **37** (2007), 959–976.

[6] N. Alon, E. Fischer, I. Newman, and A. Shapira, A combinatorial characterization of the testable graph properties: it's all about regularity. *SIAM J. Comput.* **39** (2009), 143–167.

[7] N. Alon and J. Fox, Easily testable graph properties. *Combin. Probab. Comput.* **24** (2015), 646–657.

[8] N. Alon and A. Shapira, On an extremal hypergraph problem of Brown, Erdős and Sós. *Combinatorica* **26** (2006), 627–645.

[9] N. Alon and A. Shapira, A characterization of the (natural) graph properties testable with one-sided error. *SIAM J. Comput.* **37** (2008), 1703–1727.

[10] N. Alon and A. Shapira, A separation theorem in property testing. *Combinatorica* **28** (2008), 261–281.

[11] N. Alon, A. Shapira, and B. Sudakov, Additive approximation for edge-deletion problems. *Ann. of Math.* **170** (2009), 371–411.

[12] N. Alon and C. Shikhelman, Many T copies in H-free graphs. *J. Combin. Theory Ser. B* **121** (2016), 146–172.

[13] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, Proof verification and the hardness of approximation problems. *J. ACM* **45** (1998), 501–555.

[14] S. Arora and S. Safra, Probabilistic checking of proofs: a new characterization of NP. *J. ACM* **45** (1998), 70–122.

[15] T. Austin and T. Tao, Testability and repair of hereditary hypergraph properties. *Random Structures Algorithms* **36** (2010), 373–463.

[16] F. A. Behrend, On sets of integers which contain no three terms in arithmetic progression. *Proc. Natl. Acad. Sci. USA* **32** (1946), 331–332.

[17] A. Bhattacharyya, E. Grigorescu, and A. Shapira, A unified framework for testing linear-invariant properties. *Random Structures Algorithms* **46** (2015), 232–260.

[18] M. Blum, M. Luby, and R. Rubinfeld, Self-testing/correcting with applications to numerical problems. *J. Comput. System Sci.* **47** (1993), 549–595.

[19] B. Bollobás and E. Györi, Pentagons vs. triangles. *Discrete Math.* **308** (2008), 4332–4336.

[20] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy, and K. Vesztergombi, Graph limits and parameter testing. In *Proc. of STOC 2006*, pp. 261–270, 2006.

[21] W. G. Brown, P. Erdős, and V. T. Sós, On the existence of triangulated spheres in 3-graphs and related problems. *Period. Math. Hungar.* **3** (1973), 221–228.

[22] W. G. Brown, P. Erdős, and V. T. Sós, Some extremal problems on r-graphs. In *New Directions in the Theory of Graphs, Proc. 3rd Ann Arbor Conference on Graph Theory*, pp. 55–63, Academic Press, New York, 1973.

[23] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs. *Combinatorica* **9** (1989), 345–362.

[24] D. Conlon and J. Fox, Bounds for graph regularity and removal lemmas. *Geom. Funct. Anal.* **22** (2012), 1191–1256.

[25] D. Conlon and J. Fox, Graph removal lemmas. In *Surveys in Combinatorics*, pp. 1–50, Cambridge University Press, 2013.

[26] D. Conlon, L. Gishboliner, Y. Levanzov, and A. Shapira, A new bound for the Brown–Erdős–Sós problem, submitted.

[27] R. Duke, H. Lefmann, and V. Rödl, A fast approximation algorithm for computing the frequencies of subgraphs in a given graph. *SIAM J. Comput.* **24** (1995), 598–620.

[28] P. Erdős, Problems and results in combinatorial number theory. In *Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, Bordeaux, 1974)*, pp. 295–310, Astérisque 24–25, Soc. Math. France, Paris, 1975.

[29] P. Erdős, P. Frankl, and V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.* **2** (1986), 113–121.

[30] E. Fischer and I. Newman, Testing versus estimation of graph properties. *SIAM J. Comput.* **37** (2007), 482–501.

[31] J. Fox, A new proof of the graph removal lemma. *Ann. of Math.* **174** (2011), 561–579.

[32] J. Fox, L. Gishboliner, A. Shapira, and R. Yuster, The Removal Lemma for Tournaments. *J. Combin. Theory Ser. B* **136** (2019), 110–134.

[33] J. Fox and L. M. Lovász, A tight bound for Green's arithmetic triangle removal lemma in vector spaces. *Adv. Math.* **321** (2017), 287–297.

[34] J. Fox and L. M. Lovász, A tight lower bound for Szemerédi's regularity lemma. *Combinatorica* **37** (2017), 911–951.

[35] J. Fox, J. Pach, and A. Suk, A polynomial regularity lemma for semi-algebraic hypergraphs andits applications in geometry and property testing. *SIAM J. Comput.* **45**, 2199–2223.

[36] P. Frankl and V. Rödl, Extremal problems on set systems. *Random Structures Algorithms* **20** (2002), 131–164.

[37] A. Frieze and R. Kannan, Quick approximation to matrices and applications. *Combinatorica* **19** (1999), 175–220.

[38] Z. Füredi and M. Ruszinkó, Uniform hypergraphs containing no grids. *Adv. Math.* **240** (2013), 302–324.

[39] H. Furstenberg, Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31** (1997), 204–256.

[40] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations. *J. Anal. Math.* **34** (1978), 275–291.

[41] L. Gishboliner, Y. Levanzov, and A. Shapira, An approximate version of the Gowers–Long conjecture, submitted.

[42] L. Gishboliner and A. Shapira, Efficient removal without efficient regularity. *Combinatorica* **39** (2019), 639–658.

[43] L. Gishboliner and A. Shapira, A generalized Turán problem and its applications. *Int. Math. Res. Not.* **11** (2020), 3417–3452.

[44] L. Gishboliner and A. Shapira, Constructing dense grid-free linear 3-graphs. *Proc. Amer. Math. Soc.*, to appear.

[45] L. Gishboliner and A. Shapira, Removal lemmas with polynomial bounds. *Int. Math. Res. Not.*, to appear.

[46] L. Gishboliner and A. Shapira, Testing graphs against an unknown distribution. *Israel J. Math.*, to appear.

[47] L. Gishboliner, A. Shapira, and H. Stagni, Testing Linear Inequalities of Subgraph Statistics. *Random Structures Algorithms* **58** (2021), 468–479.

[48] O. Goldreich, *Introduction to Property Testing*. Cambridge University Press, 2017.

[49] O. Goldreich, Testing graphs in vertex-distribution-free models. In *Proc. STOC 2016*, pp. 527–534, 2016.

[50] O. Goldreich, S. Goldwasser, and D. Ron, Property testing and its connection to learning and approximation. *J. ACM* **45** (1998), 653–750.

[51] O. Goldreich and I. Shinkar, Two-sided error proximity oblivious testing. *Random Structures Algorithms* **48** (2016), 341–383.

[52] W. T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma. *Geom. Funct. Anal.* **7** (1997), 322–337.

[53] W. T. Gowers, A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), 465–588.

[54] W. T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math.* **166** (2007), 897–946.

[55] W. T. Gowers and J. Long, The length of an s-increasing sequence of r-tuples. *Combin. Probab. Comput.*, to appear.

[56] B. Green, A Szemerédi-type regularity lemma in abelian groups, with applications. *Geom. Funct. Anal.* **15** (2005), 340–376.

[57] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions. *Ann. of Math.* **167** (2008), 481–547.

[58] E. Györi and H. Li, The maximum number of triangles in C_{2k+1}-free graphs. *Combin. Probab. Comput.* **21** (2012), 187–191.

[59] A. Hassidim, J. Kelner, H. Nguyen, and K. Onak, Local graph partitions for approximation and testing. In *50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 22–31, 2009.

[60] C. Hoppen, Y. Kohayakawa, R. Lang, H. Lefmann, and H. Stagni, On the query complexity of estimating the distance to hereditary graph properties. *SIAM J. Discrete Math.* **35** (2021), 1238–1251.

[61] F. Joos, J. Kim, D. Kuhn, and D. Osthus, A characterization of testable hypergraph properties. In *IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* pp. 859–867, 2017.

[62] S. Kalyanasundaram and A. Shapira, A wowzer type lower bound for the strong regularity lemma. *Proc. Lond. Math. Soc.* **106** (2013), 621–649.

[63] T. Kaufman and M. Sudan, Algebraic property testing: the role of invariance. In *Proc. 40th Annual ACM Symposium on the Theory of Computing (STOC)*, pp. 403–412, 2008.

[64] P. Keevash and J. Long, The Brown–Erdős–Sós Conjecture for hypergraphs of large uniformity. *Proc. Amer. Math. Soc.*, to appear.

[65] D. Král', O. Serra, and L. Vena, A combinatorial proof of the removal lemma for groups. *J. Combin. Theory Ser. A* **116** (2009), 971–978.

[66] D. Král', O. Serra, and L. Vena, A removal lemma for systems of linear equations over finite fields. *Israel J. Math.* **187** (2012), 193–207.

[67] A. Lev-Ran and A. Shapira, A lower bound for the strong cylinder lemma, in preparation.

[68] L. Lovász, *Large networks and graph limits*. AMS, 2012.

[69] L. Lovász and B. Szegedy, Szemerédi's lemma for the analyst. *Geom. Funct. Anal.* **17** (2007), 252–270.

[70] G. Moshkovitz and A. Shapira, A short proof of Gowers' lower bound for the regularity lemma. *Combinatorica* **36** (2016), 187–194.

[71] G. Moshkovitz and A. Shapira, A sparse regular approximation lemma. *Trans. Amer. Math. Soc.* **371** (2019), 6779–6814.

[72] G. Moshkovitz and A. Shapira, A tight bound for hypergraph regularity. *Geom. Funct. Anal.* **29** (2019), 1531–1578.

[73] B. Nagle, V. Rödl, and M. Schacht, The counting lemma for regular k-uniform hypergraphs. *Random Structures Algorithms* **28** (2006), 113–179.

[74] R. Nenadov, B. Sudakov, and M. Tyomkyn, Proof of the Brown–Erdős–Sós conjecture in groups. *Math. Proc. Cambridge Philos. Soc.* **169** (2020), 323–333.

[75] V. Rödl, Quasi-randomness and the regularity method in hypergraphs. In *Proceedings of the International Congress of Mathematicians (ICM) 1*, pp. 571–599, 2015.

[76] V. Rödl and R. Duke, On graphs with small subgraphs of large chromatic number. *Graphs Combin.* **1** (1985), 91–96.

[77] V. Rödl and M. Schacht, Regular partitions of hypergraphs: counting lemmas. *Combin. Probab. Comput.* **16** (2007), 887–901.

[78] V. Rödl and M. Schacht, Regular partitions of hypergraphs: regularity lemmas. *Combin. Probab. Comput.* **16** (2007), 833–885.

[79] V. Rödl and M. Schacht, Generalizations of the removal lemma. *Combinatorica* **29** (2009), 467–501.

[80] V. Rödl and M. Schacht, Regularity lemmas for graphs. In *Fete of Combinatorics and Computer Science*, pp. 287–325, Bolyai Soc. Math. Stud. 20, 2010.

[81] V. Rödl and J. Skokan, Regularity lemma for k-uniform hypergraphs. *Random Structures Algorithms* **25** (2004), 1–42.

[82] K. F. Roth, On certain sets of integers (II). *J. Lond. Math. Soc.* **29** (1954), 20–26.

[83] R. Rubinfeld and M. Sudan, Robust characterizations of polynomials with applications to program testing. *SIAM J. Comput.* **25** (1996), 252–271.

[84] I. Ruzsa, Solving a linear equation in a set of integers I. *Acta Arith.* **65** (1993), 259–282.

[85] I. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles. In *Combinatorics (Keszthely, 1976), Volume II*, pp. 939–945, Colloq. Math. Soc. János Bolyai 18, North Holland, Amsterdam, 1978.

[86] S. Sapir and A. Shapira, The induced removal lemma in sparse graphs. *Combin. Probab. Comput.* **29** (2020), 153–162.

[87] G. N. Sárközy and S. Selkow, An extension of the Ruzsa–Szemerédi theorem. *Combinatorica* **25** (2004), 77–84.

[88] C. Shangguan and I. Tamo, Sparse hypergraphs with applications to coding theory. *SIAM J. Discrete Math.* **34** (2020), 1493–1504.

[89] A. Shapira, A proof of Green's conjecture regarding the removal properties of sets of linear equations. *J. Lond. Math. Soc.* **81** (2010), 355–373.

[90] A. Shapira and M. Tyomkyn, A Ramsey variant of the Brown–Erdős–Sós conjecture. *Bull. Lond. Math. Soc.*

[91] D. Solymosi and J. Solymosi, Small cores in 3-uniform hypergraphs. *J. Combin. Theory Ser. B* **122** (2017), 897–910.

[92] J. Solymosi, A note on a question of Erdős and Graham. *Combin. Probab. Comput.* **13** (2004), 263–267.

[93] E. Szemerédi, On sets of integers containing no four elements in arithmetic progression. *Acta Math. Acad. Sci. Hung.* **20** (1969), 89–104.

[94] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27** (1975), 199–245.

[95] E. Szemerédi, Regular partitions of graphs. In *Proc. Colloque Inter. CNRS*, pp. 399–401, 1978.

[96] T. Tao, The dichotomy between structure and randomness, arithmetic progressions, and the primes. In *Proc. Intern. Congress of Math. I*, pp. 581-–608, Eur. Math. Soc., Zurich, 2006.

[97] T. Tao, A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A* **113** (2006), 1257–1280.

[98] A. Thomason, Pseudorandom graphs. In *Random graphs '85 (Poznań, 1985)*, pp. 307–331, North-Holl. Math. Stud. 144, North-Holland, Amsterdam, 1987.

[99] J. Tidor and Y. Zhao, Testing linear-invariant properties. In *IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 1180–1190, 2020.

[100] P. Turán, On an extremal problem in graph theory. *Mat. Fiz. Lapok* **48** (1941), 436–452.

ASAF SHAPIRA

School of Mathematical Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel, asafico@tauex.tau.ac.il